
## SENIOR SIX TERM 2

# TOPIC 3/5: Sampling Distribution

**Competency:** The learner estimates population parameters by using sample distributions to predict and eliminate variability in research and collection of statistical data.

**NB. You need prior knowledge of normal distribution**

## Distribution of sample mean of a normal distribution population

If a random variable X of a sample of size n from a normal distribution with mean μ and variance $\sigma^2$, then distribution of the sample mean $\bar{x}$ is also said to be normally distributed with mean μ and variance $\frac{\sigma^2}{n}$, such that $\bar{x} \approx \left(\mu,\ \frac{\sigma^2}{n}\right)$

Then $Z = \frac{\bar{x}-\mu}{\frac{\sigma^2}{n}}$

**Example 1**

At a certain school, the masses of students are normally distributed with mean 70kg and standard deviation 5kg. If 4 students are randomly selected, find the probability that their mean is less than 65.

P$(\bar{X} < 65) = P\left(Z < \frac{65-70}{\frac{5}{\sqrt{4}}}\right) = P(Z < -2)$

$P(Z < -2) = P(Z > 2) = 0.5 - P(0 < Z < 2)$ =0.5 – 0.4772 = 0.0228

**Example 2**

A random sample of size 15 is taken from a normal population with mean 60 and standard deviation 4. Find the probability that the mean of the sample is less than 58

P$(\bar{X} < 58) = P\left(Z < \frac{58-60}{\frac{4}{\sqrt{15}}}\right) = P(Z < -1.936)$

$P(Z < -1.936) = P(Z > 1.936) = 0.5 - P(0 < Z < 1.936)$ =0.5 – 0.4736 = 0.0264

**Example 3**

The height of students are normally distributed with mean 164cm and standard deviation 7.2cm. Calculate the probability that the mean height of a sample of 36 students will be between 162cm and 166cm.

$$P(162 < \overline{X} < 166) = P\left(\frac{162-164}{\frac{7.2}{\sqrt{36}}} < Z < \frac{166-164}{\frac{7.2}{\sqrt{36}}}\right) = P(-1.667 < Z < 1.667)$$

$P(-1.667 < Z < 1.667) = 2\ x\ P(0 < Z < 1.667$ = 2 x 0.4522 = 0.9044

### Example 4

The height of a certain plant follows a normal distribution with mean 21cm and standard deviation $\sqrt{90}cm$. A random sample of 10 plants is taken and the mean height calculated. Find the probability that this sample mean lies between 18cm and 27 cm

$$P(18 < \overline{X} < 27) = P\left(\frac{18-21}{\frac{\sqrt{90}}{\sqrt{10}}} < Z < \frac{27-21}{\frac{\sqrt{90}}{\sqrt{10}}}\right) = P(-1 < Z < 2)$$

$P(-1 < Z < 2) = P(0 < Z < 1) + P(0 < Z < 2)$ = 0.3413 + 0.4772 = 0.8185

### Example 5

A large number of random sample of size n is taken from a distribution X where X ~N(74, 36) and the sample mean $\overline{x}$ for each sample is noted. If $P(\overline{x} > 72) = 0.854, find\ the\ value\ of\ n.$

$$P(\overline{X} > 72) = P\left(Z > \frac{72-74}{\frac{6}{\sqrt{n}}}\right) = 0.854$$

$$P\left(Z > \frac{-\sqrt{n}}{3}\right) = 0.854$$

From table Z = - 1.054

$$\frac{-\sqrt{n}}{3} = -1.054$$

n = 10

### Example 6

The distribution of a random variable x is X ~N(25, 340) and the sample mean $\overline{x}$ for each sample is calculated. If $P(\overline{x} > 28) = 0.005, find\ the\ value\ of\ n.$

$$P(\overline{X} > 28) = P\left(Z > \frac{28-25}{\frac{\sqrt{340}}{\sqrt{n}}}\right) = 0.005$$

$$P\left(Z > \frac{3\sqrt{n}}{\sqrt{340}}\right) = 0.005$$

From table Z = 2.576

$\frac{3\sqrt{n}}{\sqrt{340}} = 2.576$

n = 250

## Revision exercise 1

1. If X~$N$(200, 80) and a random sample of size 5 is taken from the distribution, find the probability that the sample mean
   (i) is greater than 207 [0.0401]
   (ii) lies between 201 and 209 [0.3891]
2. If X~$N$(200, 10) and a random sample of size 10 is taken from the distribution, find the probability that the sample mean lies outside the range 198 and 205 [0.3206]
3. If X~$N$(50, 12) and a random sample of size 12 is taken from the distribution, find the probability that the sample mean
   (i) Is less than48.5 [0.0668]
   (ii) Is less than 52.3[0.9893]
   (iii) Lie between 50.7 and 51.7 [0.1974]
4. Biscuits are produced with weight (W g) where W is N(10, 4) and are packed at random into boxes consisting of 25 biscuits. Find the probability that
   (i) a biscuit chosen at random weigh between 9.25g and 10.7g [0.2924]
   (ii) the content of a box weighs between 245g and 255g [0.0796]
   (iii) the average weight of the biscuit in the box lies between 9.7g and 10.3g [0.5468]
5. A normal distribution has a mean of 40 and standard deviation of 4. If 25 items are drawn at random, find the probability that their mean
   (i) 41.4 or more [0.0401]
   (ii) Between 38.7 and 40 .7 [0.7571]
   (iii) Less than 39.5 [0.2660]
6. A random sample of size 25 is taken from a normal population with mean 60 and standard deviation 4. Find the probability that the mean of the sample
   (i) Less than 58 [0.0062]
   (ii) Greater than 58 [0.9918]
   (iii) Between 58 and 62[0.9876]
7. At St. Noa Junior, the marks of the pupils can be modelled by a normal distribution with mean 70% and standard deviation 5%. If four pupils are chosen at random, find the probability that the mean mark is
   (i) Less than 65% [0.9772]
   (ii) Greater than 65% [0.0228]
   (iii) Greater than 75%[0.0228]
   (iv) Between 72% and 75% [0.1891]
8. The volume of soda in bottle are normally distributed with mean 758ml and standard deviation of 12ml. a random sample of 10 bottles is taken and mean volume is found. Find the probability that the sample mean is less than 750ml. [0.0176]
9. The height of cassava plants are normally distributed with mean of 2m and standard deviation of 40cm. a random sample of 50 cassava plants is taken and the mean height found. Find the probability that the sample mean lies between 195cm and 205cm. [0.6234]

10. In an examination, marks are normally distributed with mean 64.5 and variance 64. The mean ark in a random sample of 100 scripts is denoted by X. find
    (i) P(X> 65.5) [0.1056]
    (ii) P(63.8 < $\overline{X}$ < 64.5[0.3092]
11. The marks of an examination were normally distributed. 20% of the students scored below 40 marks while 10%of the students scored above 75 marks
    (i) Find the mean mark and standard deviation of the students [μ = 53.87, σ = 16.473]
    (ii) If 25 students were chosen at random from those who sat for the examination, what is the probability that their average mark exceeds 60. [0.0313]
    (iii) If a sample of 8 students were chosen, find the probability that not more than 3 scored between 45 and 65 marks. [0.5419]
12. The life time of batteries produced by a certain factory is normally distributed. Out of 10,000 batteries selected at random, 668have life tie less than 130 hours and 228 have life time more than 200 hours.
    (i) Find the mean mark and standard deviation of the battery life time[μ =160, σ = 20]
    (ii) Find the percentage of the batteries with life time between 150 and 180 hours.
    (iii) If the sample of 25 batteries is selected at random, find the probability that the mean of the life time exceeds 165 hours [0.1056]
13. A normal distribution has a mean of 30 and a variance of 5. Find the probability that
    (i) The average of 10 observation exceeds 30.5 [0.2399]
    (ii) The average of 40 observation exceeds 30.5[0.0787]
    (iii) The average of 100 observation exceeds 30.5 [0.0127]
    (iv) Find n such that the probability that the average of observations exceed 30.5 is less than 1%, [n > 108]
14. The random variable is such that X$\sim N$(μ, 4). A random sample size n is taken from the population. Find the least n such that P($\left|\overline{X} - \mu\right| < 0.5) = 0.95$ [62]
15. Boxes made in a factory have weight which are normally distributed with a mean of 4.5kg and a standard deviation of 2.0kg. if a sample of 16 boxes is drawn at random, find the probability that their mean is
    (i) between 4.6 and 4.7 kg [0.0761]
    (ii) between 4.3 and 4.7g [0.3108]
16. the masses of soap powder in a certain packet are normally distributed with mean 842g and variance 225g. find the probability that a random sample of 120 packets has sample mean mass
    (i) between 844g and 846g [0.0702]
    (ii) less than 843g[0.7673]

## Estimation of population parameters

Statistical estimation is used to describe the unknown characteristics of the population (population parameters) by using sample characteristics.

A sample is a representation of the population parameter such as population mean, μ and population variance, $\sigma^2$.

**Types of parameter estimation**

- point estimation
- interval estimation

**(a) point estimates**

(i) the unbiased estimate of the population mean, μ is

$\overline{x} = \frac{\sum x}{n}$ $or$ $\overline{x} = \frac{\sum fx}{\sum f}$ where $\overline{x}$ is sample mean

(ii) the unbiased estimate of the population variance, $\sigma^2$ is $\hat{\sigma}^2$ where $\hat{\sigma}^2 = \frac{n}{n-1}s^2$ where $s^2$ is sample variance

OR $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2\right]$ or $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

**Example 7**

Find the best unbiased estimate of mean μ and variance σ2 of the population from each of the following sample is drawn

(i) 46, 48, 50, 45, 53, 50, 48, 51

Solution

| x | f | fx | $fx^2$ |
|---|---|---|---|
| 45 | 1 | 45 | 2025 |
| 46 | 1 | 46 | 2116 |
| 48 | 2 | 96 | 4608 |
| 50 | 2 | 100 | 5000 |
| 51 | 1 | 51 | 2601 |
| 53 | 1 | 53 | 2809 |
| | $\sum f$= 8 | $\sum fx$= 391 | $\sum fx^2$= 19159 |

Unbiased estimate for the mean $\overline{x} = \frac{\sum fx}{\sum f} = \frac{391}{8} = 48.875$

The unbiased estimate of the population variance, $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

$= \frac{8}{8-1}\left[\frac{19159}{8} - \left(\frac{391}{8}\right)^2\right]$ = 6.982

(ii) $\sum x$ = 100, $\sum x^2$ = 1028, n = 10

Unbiased estimate for the mean $\overline{x} = \frac{\sum fx}{\sum f} = \frac{100}{10} = 10$

The unbiased estimate of the population variance, $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

$$= \frac{10}{10-1}\left[\frac{1028}{10} - \left(\frac{100}{10}\right)^2\right] = 3.11$$

(iii) $\sum x = 120, \sum x^2 = 2102, n = 8$

Unbiased estimate for the mean $\overline{x} = \frac{\sum fx}{\sum f} = \frac{120}{8} = 15$

The unbiased estimate of the population variance, $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

$$= \frac{8}{8-1}\left[\frac{2102}{8} - \left(\frac{120}{8}\right)^2\right] = 43.14$$

(iv) $\sum x = 330, \sum x^2 = 23700, n = 34$

Unbiased estimate for the mean $\overline{x} = \frac{\sum fx}{\sum f} = \frac{330}{34} = 9.71$

The unbiased estimate of the population variance, $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

$$= \frac{34}{34-1}\left[\frac{23700}{34} - \left(\frac{330}{34}\right)^2\right] = 621.12$$

(v) $\sum x = 738, \sum x^2 = 16526, n = 50$

Unbiased estimate for the mean $\overline{x} = \frac{\sum fx}{\sum f} = \frac{738}{50} = 14.76$

The unbiased estimate of the population variance, $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

$$= \frac{50}{50-1}\left[\frac{16526}{50} - \left(\frac{738}{50}\right)^2\right] = 114.96$$

**Example 8**

The fuel consumption of a new car model was being tested. In one trials 8 cars chosen at random were driven under identical conditions and distance x km covered on one litre of petro was recorded. The following results were obtained. $\sum x = 152.98, \sum x^2 = 2927.1$. Calculate the unbiased estimate of the mean and variance of the distance covered by the car.

Solution

Unbiased estimate for the mean $\overline{x} = \frac{\sum fx}{\sum f} = \frac{152.98}{8} = 19.1225$

The unbiased estimate of the population variance, $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

$= \frac{8}{8-1}\left[\frac{2927.1}{8} - \left(\frac{152.98}{8}\right)^2\right] = 0.25$

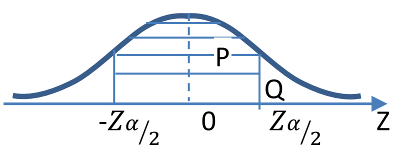
**(b) Interval estimate**

Here we are interested in obtaining the interval over which the true population mean lies (confidence interval)

The unbiased estimate of the population mean, μ is $\overline{x}$

Z = $\frac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ where n is the sample size

Z is the area under the normal curve leaving an area of $\frac{\alpha}{2}$ on either side of the curve




$P\left(-Z\alpha/_2 < Z < Z\alpha/_2\right) = \frac{\alpha}{2}$

$P\left(-Z\alpha/_2 < \frac{\overline{x}-\mu}{\frac{\sigma}{\sqrt{n}}} < Z\alpha/_2\right) = \frac{\alpha}{2}$

$P\left(\overline{x} - Z\alpha/_2\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + Z\alpha/_2\frac{\sigma}{\sqrt{n}}\right) = \frac{\alpha}{2}$

Confidence interval $\left[\overline{x} - Z\alpha/_2\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + Z\alpha/_2\frac{\sigma}{\sqrt{n}}\right]$

Confidence Limits $\left[\overline{x} - Z\alpha/_2\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + Z\alpha/_2\frac{\sigma}{\sqrt{n}}\right]$

Or $\mu < \overline{x} \pm Z\alpha/_2\frac{\sigma}{\sqrt{n}}$

**(i) Confidence interval for population mean μ**

- of a normal or non-normal population
- with known population variance σ2 or standard deviation σ
- using any sample size

The confidence interval is obtained from $\mu < \overline{x} \pm Z\alpha/_2\frac{\sigma}{\sqrt{n}}$ where $\overline{x}$ is sample mean

**Example 9**

The length of a bar of a metal is normally distributed with mean of 115cm and standard deviation of 3cm. find the 95% confidence limits for the length of the bar

$\frac{\alpha}{2} = \frac{0.95}{2} = 0.475$

$Z\alpha/_2 = 1.96$

$\mu < \overline{x} \pm Z\alpha/_2\frac{\sigma}{\sqrt{n}}$

$\mu < 115 \pm 1.96\frac{3}{\sqrt{1}}$

Lower limit =109.12

Upper limit = 120.88

**Example 10**

The mass of vitamin in a capsule is normally distributed with standard deviation 0.042mg. a random sample of 5 capsules was taken and the mean mass of vitamin e found to be 5.12. calculate a symmetric confidence interval for the population mean mass.

$\frac{\alpha}{2} = \frac{0.95}{2} = 0.475$

$Z\alpha/_2 = 1.96$

$\mu < \overline{x} \pm Z\alpha/_2\frac{\sigma}{\sqrt{n}}$

$\mu < 5.12 \pm 1.96\frac{0.042}{\sqrt{5}}$

Lower limit =5.08

Upper limit = 5.16

**Example 11**

It is known that an examination paper is marked in such a way that the standard deviation of the marks is 15.1. in a certain school, 80 candidates took the examination and they had an average mark of 57.4. find

(i) 95% confidence limits for the mean mark in the exam.

$\frac{\alpha}{2} = \frac{0.95}{2} = 0.475$

$Z\alpha/_2 = 1.96$

$\mu < \bar{x} \pm Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$

$\mu < 5.12 \pm 1.96 \frac{15.1}{\sqrt{80}}$

Lower limit =54.091

Upper limit = 60.709

(ii) 99% confidence limits for the mean mark in the exam.

$\frac{\alpha}{2} = \frac{0.99}{2} = 0.495$

$Z\alpha/_2 = 2.575$

$\mu < \bar{x} \pm Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$

$\mu < 5.12 \pm 2.595 \frac{15.1}{\sqrt{80}}$

Lower limit =53.053

Upper limit = 61.746

**Example 12**

After a particular rainy night, 12 worms were picked and their length in cm measured;

9.5, 9.5, 11.2, 10.6, 9.9, 11.1, 10.9, 9.8, 10.1, 10.2, 10.9, 11.0

Assuming that this sample came from a normal population with standard deviation 2, find the 95 confidence interval for the mean length of all the worms

$$\bar{x} = \frac{\sum x}{n} = \frac{9.5 + 9.5 + 11.2 + 10.6 + 9.9 + 11.1 + 10.9 + 9.8 + 10.1 + 10.2 + 10.9 + 11.0}{12}$$

= 10.39

$\frac{\alpha}{2} = \frac{0.95}{2} = 0.475$

$Z\alpha/_2 = 1.96$

$\mu < \bar{x} \pm Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$

$\mu < 10.39 \pm 1.96 \frac{2}{\sqrt{12}}$

Lower limit =9.81

Upper limit = 10.97

The height of students are normally distributed with mean µ and standard deviation σ. On the bsis of results obtained from a random sample of 100 students from school, the 95% confidence interval of the mean was calculated and found to be (177.22cm, 179.18cm). Calculate

(i) the value of the sample mean

$\frac{\alpha}{2} = \frac{0.95}{2} = 0.475$

$Z\alpha/_2 = 1.96$

$\mu < \bar{x} \pm Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$

$177.22 < \bar{x} - 1.96 \frac{\sigma}{\sqrt{100}}$ .... (i)

$179.18 < \bar{x} + 1.96 \frac{\sigma}{\sqrt{100}}$ .... (ii)

Eqn. (i) + eqn. (ii)

$2\bar{x}$ = 356.4; $\bar{x}$ = 178.2

(ii) the value of standard deviation

$177.22 < 178.2 - 1.96\frac{\sigma}{\sqrt{100}}$;

σ = 5

(iii) 90% confidence interval of the mean, μ

$\frac{\alpha}{2} = \frac{0.90}{2} = 0.45; Z\alpha/_2 = 1.645$

**Example 13**

A plant produces steel sheets whose weight are normally distributed with standard deviation of 2.4kg. A random sample of 36 sheets had a mean weight of 31.4kg.

(i) Find the 99% confidence limit for the population

$\frac{\alpha}{2} = \frac{0.99}{2} = 0.495$

$Z\alpha/_2 = 2.575$

$\mu < \overline{x} \pm Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$

$\mu < 31.4 \pm 2.595\frac{2.4}{\sqrt{36}}$

Lower limit =30.37kg

Upper limit = 32.43kg

(ii) Find the width of the 99% confidence limit

=32.43kg - 30.37kg = 2.06kg

**Example 47**

The marks scored by students are normally distributed with mean μ and standard deviation1.3. it is required to have 95% confidence interval for μ with width less than 2. Find the least number of students that be sampled to achieve this.

$\frac{\alpha}{2} = \frac{0.95}{2} = 0.475$

$Z\alpha/_2 = 1.96$

width = $2\ xZ\alpha/_2 \frac{\sigma}{\sqrt{n}}$ < 2

$2\ x\ 1.96\frac{\sigma}{\sqrt{n}}$ < 2

$2\ x\ 1.96\frac{\sigma}{2}$ < $\sqrt{n}$; n < 6.49

n > 6.49

the least number = 7

**(ii) Confidence interval for population mean μ**

- of a normal or non-normal population
- with unknown population variance σ2 or standard deviation σ
- using a large sample size (n ≥ 30)

If the population variance σ2 s not given or unknown, the confidence interval is obtained from $\mu < \overline{x} \pm Z\alpha/_2 \frac{\hat{\sigma}}{\sqrt{n}}$ where $\overline{x}$ is sample mean, $\hat{\sigma} = \frac{n}{n-1}s^2$ and s = sample variance

$\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2\right]$

Or

$\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

### Example 14

the fuel consumption of a new car was being tested. In one trials 50 cars chosen at random were driven under identical conditions and the distance x km covered on one litre of petrol was recorded. the following results were obtained. $\sum x = 525$ , $\sum x^2 = 5625$. Calculate the 95%confidence interval for the mean petrol consumption, in km per litre of cars of this type..

Unbiased estimate for the mean $\overline{x} = \frac{\sum fx}{\sum f} = \frac{525}{50} = 10.5$

The unbiased estimate of the population variance, $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

$= \frac{50}{50-1}\left[\frac{5625}{50} - \left(\frac{525}{50}\right)^2\right]$ = 2.2952

$\frac{\alpha}{2} = \frac{0.95}{2} = 0.475$

$Z\alpha/_2 = 1.96$

$\mu < \overline{x} \pm Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$

$\mu < 10.5 \pm 1.96 \frac{\sqrt{2.2952}}{\sqrt{50}}$

Lower limit =10.08km/litre

Upper limit = 10.92km/litre

### Example 15

The height x cm of each man in a random sample of 200 men living in Nairobi was measured. The following results were obtained $\sum x = 35050$ , $\sum x^2 = 6163109$.

(a) calculate the unbiased estimate of the mean and variance of the heights of men living Nairobi

Unbiased estimate for the mean $\overline{x} = \frac{\sum fx}{\sum f} = \frac{35050}{200} = 175.25$

The unbiased estimate of the population variance, $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

$= \frac{200}{200-1}\left[\frac{6163109}{200} - \left(\frac{35050}{200}\right)^2\right]$ = 103.5

(b) Determine the 90% confidence interval for the mean height of mean living in Nairobi.

$\frac{\alpha}{2} = \frac{0.90}{2} = 0.45$

$Z\alpha/_2 = 1.645$

$\mu < \overline{x} \pm Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$

$\mu < 175.25 \pm 1.645 \frac{\sqrt{103.5}}{\sqrt{200}}$

Lower limit =174.07cm

Upper limit = 176.43cm

### Example 16

A random sample of 100 observations from a normal population with mean μ gave the following results $\sum x = 8200$ , $\sum x^2 = 686000$.

(a) calculate the unbiased estimate of the mean and variance of the heights of men living Nairobi

Unbiased estimate for the mean $\overline{x} = \frac{\sum fx}{\sum f} = \frac{8200}{100} = 82$

The unbiased estimate of the population variance, $\hat{\sigma}^2 = \frac{n}{n-1}\left[\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2\right]$

$= \frac{200}{200-1}\left[\frac{68600}{100} - \left(\frac{8200}{100}\right)^2\right] = 11.72$

(b) Determine the 98% confidence interval for the mean

| | | |
|---|---|---|
| $\frac{\alpha}{2} = \frac{0.98}{2} = 0.49$ | $\mu < \overline{x} \pm Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$ | Lower limit =79.274 |
| $Z\alpha/_2 = 2.326$ | $\mu < 82 \pm 2.326 \frac{11.72}{\sqrt{100}}$ | Upper limit = 84.726 |

(c) determine the 99% confidence interval for the mean

| | | |
|---|---|---|
| $\frac{\alpha}{2} = \frac{0.99}{2} = 0.495$ | $\mu < \overline{x} \pm Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$ | Lower limit =78.981 |
| $Z\alpha/_2 = 2.575$ | $\mu < 82 \pm 2.595 \frac{11.72}{\sqrt{100}}$ | Upper limit = 85.726 |

**Example 17**

The mean and standard deviation of a random sample of size 100 is 900 and 60 respectively. Given that the population is normally distributed, find 96% confidence interval of the population mean.

$$\hat{\sigma} = \sqrt{\frac{n}{n-1}s^2} = \sqrt{\frac{100}{100-1}x\ 60^2} = 60.302$$

| | | |
|---|---|---|
| $\frac{\alpha}{2} = \frac{0.96}{2} = 0.48$ | $\mu < \overline{x} \pm Z\alpha/_2 \frac{\sigma}{\sqrt{n}}$ | Lower limit =887.614 |
| $Z\alpha/_2 = 2.054$ | $\mu < 900 \pm 2.054 \frac{60.302}{\sqrt{100}}$ | Upper limit = 912.386 |

**Revision exercise 2**

1. the concentration in mg per litre of a trace element in 7 randomly chosen samples of water from a spring were 240.8, 237.3, 236.7, 236.6, 234.2, 233.9, 232.5.
   Determine the unbiased mean and variance of the concentration of the trace element per litre from spring [236, 7.58]

2. Cartons of oranges are filled by a machine. A sample of 10 cartons selected at random from the population contained the following quantities in ml) 201.2, 205.0, 209.1, 202.3,204.6, 206.4, 210.1,201.9, 203.7, 207.3. Determine the unbiased mean and variance of the population from which the sample was taken. [203.16, 9.223]
3. A factory produces cans of meat whose masses are normally distributed with standard deviation 18g. A random sample of 25 cans is found to have the mean of 458g. find the 99% confidence interval for the population mean of a can of meat produced at the factory. [448.7, 467.3g]
4. The height of bounce of a tennis ball is normally distributed with standard deviation 2cm. A sample of 60 tennis balls is tested and the mean height of bounce is 140cm. Find
   (i) 95% [139.5, 140.51] (ii) 98% [139.4, 140.6] confidence interval for the mean height of bounce of the tennis ball
5. A random sample of 100 is taken from a population. The sample is found to have a mean of 76.0 and standard deviation of 120. Find
   (i) 90% [747.51, 748.49] (ii) 95% [747.42, 748.58] (iii) 98% [747.31, 748.69] confidence interval for the mean of the population
6. 150 bags of flour of a particular brand are weighed and the mean mass is found to be 748g with standard deviation 3.6g. Find
   (i) 90% [74.02, 77.98] (ii) 97% [73.38, 78.62] (iii) 99% [72.89, 79.11] confidence interval for the mean mass of bags of flour of this brand.
7. A random sample of 100 readings taken from a normal population gave the following data: $\overline{x}$ = 82, $\sum x^2 = 686800$. Find
   (i) 98% [79.19, 84.81] (ii) 99% [78.89, 85.11] confidence interval of the population mean
8. 80 people were asked to measure their pulse rates when they woke up in the morning. The mean was 69 beats and standard deviation 4 beats. find
   (i) 95% [68.12, 69.88] (ii) 99% [67.84, 70.16] (iii) 97% [68.0, 70.0] confidence interval of the population mean
9. The 95% confidence interval for the mean length of a particular brand of light bulbs is [1023.3h, 1101.7h]. This is interval is based on results from a sample of 36 light bulbs. Find the 99% confidence interval for the mean length of life of this brand of light bulbs assuming that the length of life is normally distributed. [1011, 1114]
10. A random sample of 6 items taken from a normal population with variance 4.5cm$^2$ gave the following data: 12.9cm, 13.2cm, 14.6cm, 12.6cm, 11.3cm, and 10.1 cm.
    (i) Find the 95% confidence interval for the population mean. [10.75, 14.15].
    (ii) What is the width of this confidence interval [3.4]
11. A random sample of 60 loaves is taken from a population whose mean masses are normally distributed with mean μ and standard deviation 10g.
    (i) calculate the width of 95% confidence interval for μ bases on the sample [5.06]
    (ii) Find the confidence level having the same width as in (i) but based on a random sample of 40 loaves. [89%]

12. The distribution of measurements of masses of a random sample of bags packed in a factory is shown below

| Mass (kg) | 72.5 | 77.5 | 82.5 | 87.5 | 92.5 | 97.5 | 102.5 | 107.5 |
|---|---|---|---|---|---|---|---|---|
| frequency | 6 | 18 | 32 | 57 | 102 | 51 | 25 | 9 |

(i) Find the mean and standard deviation of the masses [$\mu$ =91.317, $\sigma$=7.41]

(ii) find the 95% confidence limits[90.5, 92.2]

**Thank you**
**Dr. Bbosa Science**